

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Gist-brocades B.V.
(B) STREET: Wateringsweg 1
(C) CITY: Delft
(E) COUNTRY: The Netherlands
(F) POSTAL CODE (ZIP): 2611 XT

(ii) TITLE OF INVENTION: Propionibacterium Vector

(iii) NUMBER OF SEQUENCES: 13

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3555 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) ORIGINAL SOURCE:

- (A) ORGANISM: Propionibacterium freudenreichii
(C) INDIVIDUAL ISOLATE: CBS101022 IMG16545

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 273..1184
(D) OTHER INFORMATION: /gene= "ORF1"

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 1181..1438
(D) OTHER INFORMATION: /gene= "ORF2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GTGACCCCTG ACAGCCCGCG AGCAATTCAG GCGAAGATCG CACAGCTGCG CCAGGAACCTA
60

GCCGCAATGC CCGAACACGC CCCAGCCATC CCTTGGAGCA GGTGGCAGCG TCAGGGCACT
120

CGGGGGATGT TTGGCAGGGG ATCTCGAAAG AGAGTTGCT TTGCTCACAT GGCTCAACCG
180

5 GGTAACCTAAC TGATATGGGG TCTTCGTCCG CCACITTCGA CACGCCGAGG AATGGACCAC
240

GCTGAACGTG ACTCGCATGC TTCACTGCAT GT ATG GAT TCG TTC GAG ACG TTG
293

10 Met Asp Ser Phe Glu Thr Leu
1 5

TTC CCT GAG AGC TCG CTG CCA CGC AAG CCG CTG GCG TCA GCC GAG AAG
341

Phe Pro Glu Ser Trp Leu Pro Arg Lys Pro Leu Ala Ser Ala Glu Lys
10 15 20

15 TCT GGG GCG TAC CGG CAC GTG ACT CGG CAG AGG GCG CTG GAG CTG CCT
389

Ser Gly Ala Tyr Arg His Val Thr Arg Gln Arg Ala Leu Glu Leu Pro
25 30 35

20 TAC ATC GAA GCG AAC CCG TTG GTC ATG CAG TCC TTG GTC ATC ACC GAT
437

Tyr Ile Glu Ala Asn Pro Leu Val Met Gln Ser Leu Val Ile Thr Asp
40 45 50 55

CGA GAT GCT TCG GAT GCT GAC TGG GCC GCA GAC CTC GCT GGG CTG CCT
485

25 Arg Asp Ala Ser Asp Ala Asp Trp Ala Ala Asp Leu Ala Gly Leu Pro
60 65 70

TCA CCG TCC TAC GTG TCC ATG AAC CGT GTC ACG ACC ACC GGA CAC ATC
533

30 Ser Pro Ser Tyr Val Ser Met Asn Arg Val Thr Thr Thr Gly His Ile
75 80 85

GTC TAT GCC TTG AAC AAC CCT GTG TGT CTG ACC GAT GCC GCG CGG CGA
581

Val Tyr Ala Leu Lys Asn Pro Val Cys Leu Thr Asp Ala Ala Arg Arg
90 95 100

35 CGG CCT ATC AAC CTG CTC GCC CGC GTC GAG CAG GGC CTA TGC GAC GTT
629

Arg Pro Ile Asn Leu Leu Ala Arg Val Glu Gln Gly Leu Cys Asp Val
105 110 115

40 CTC GGC GGC GAT GCA TCC TAC GGG CAC CGG ATC ACA AAG AAC CCG CTC
677

Leu Gly Gly Asp Ala Ser Tyr Gly His Arg Ile Thr Lys Asn Pro Leu
120 125 130 135

AGC ACC GCC CAT GCG ACC CTC TGC GGC CCC GCA GAC GCG CTC TAC GAG
725

45 Ser Thr Ala His Ala Thr Leu Trp Gly Pro Ala Asp Ala Leu Tyr Glu
140 145 150

105290-0350260

CTG CGC CCC CTC GCA CAC ACC CTC GAC GAG ATC CAC GCA CTG CCG GAG
 773
 Leu Arg Ala Leu Ala His Thr Leu Asp Glu Ile His Ala Leu Pro Glu
 155 160 165

5 GCA GGG AAC CCG CGT CGC AAC GTC ACC CGA TCA ACG GTC GGC CGC AAC
 821
 Ala Gly Asn Pro Arg Arg Asn Val Thr Arg Ser Thr Val Gly Arg Asn
 170 175 180

10 GTC ACC CTG TTC GAC ACC ACC CGC ATG TGG GCA TAC CGG CCC GTC CCG
 869
 Val Thr Leu Phe Asp Thr Thr Arg Met Trp Ala Tyr Arg Ala Val Arg
 185 190 195

15 CAC TCC TGG GGC GGC CCG GTC GCC GAA TGG GAG CAC ACC GTA TTC GAG
 917
 His Ser Trp Gly Gly Pro Val Ala Glu Trp Glu His Thr Val Phe Glu
 200 205 210 215

20 CAC ATC CAC CTA CTG AAC GAG ACG ATC ATC GCC CAC GAA TTC GCC ACA
 965
 His Ile His Leu Leu Asn Glu Thr Ile Ile Ala Asp Glu Phe Ala Thr
 220 225 230

GGC CCC CTC GGC TTG AAC GAA CTT AAG CAC TTA TCT CGA TCC ATT TCC
 1013
 Gly Pro Leu Gly Leu Asn Glu Leu Lys His Leu Ser Arg Ser Ile Ser
 235 240 245

25 CGA TGG GTC TGG CGC AAC TTC ACC CCC GAA ACC TTC CGC GCA CGC CAG
 1061
 Arg Trp Val Trp Arg Asn Phe Thr Pro Glu Thr Phe Arg Ala Arg Gln
 250 255 260

30 AAA GCG ATC AGC CTC CGT GGA GCA TCC AAA GGC GGC AAA GAA CCC GGC
 1109
 Lys Ala Ile Ser Leu Arg Gly Ala Ser Lys Gly Gly Lys Glu Gly Gly
 265 270 275

35 CAC AAA GGC GGC ATT GCC AGT GGC GCA TCA CCG CGC GCC CAT ACC CGT
 1157
 His Lys Gly Gly Ile Ala Ser Gly Ala Ser Arg Arg Ala His Thr Arg
 280 285 290 295

40 CAA CAG TTC TTG GAG GGT CTC TCA TGACCACACG TCAACGTCTC CCCCACAACG
 1211
 Gln Gln Phe Leu Glu Gly Leu Ser
 300

GCTACACCAT CCCCCTGCT GCGAAAAAGC TCGGTGTCTC CGAGTCCACC GTCAAGCGGT
 1271

GGACTTCCGA GCCACGCCAG GAGTTCGTGG CCCGCGTTGC CGCACGCCAC GCGCGGATTC
 1331

45 GTGAGCTCCG CTCGGAGGCT CAGACCATGC GTGCGATTGC TGCCGAGGTC GGGGTTTCCG
 1391

FO5290"ES502450

TGGGCACCGT CCACTACGCG CTGAACAAGA ATCGAAGTGA CGCATGACCG TAACGCCGCA
1451

CGATGAGCAT TTCTTGATC GTGCACCGCT TGGCACTACG TTCGCGTGCG GTTGACAGT
1511

5 GCGCGCCACG TTCTTATCCT GCGGCCATTG TGGCTACAGC CAATGGGGGG CATCAGCAAC
1571

GGACGTTGAA CCGGTTGGGC AAGTGTACT CAGGCGGACA TGCCAGTCT GCGGCGCTCG
1631

10 GATTGACGGT ATGGCAGTCG TGCATGCGGC CCCACCGTCA AACTCATTCA GGTATCAGTG
1691

AGAACCCTCA TGGCACCCCC TCGTGACAGG TTCTCGTTGC GATCAGCTGC TCTCGGTGCG
1751

GGCGTGAGCG TTCTACGCT GCGCGCGCAG AAATCAGAGC TTGAGGCTGC CGGAGCGACG
1811

15 GTAGACCCGT CCGGTTGGGT GGTGCCACTG CGTGCACTCA AGGTCGTTTT TGGGGTGTC
1871

GATGAGACCT CGAATGCGCC CGGTCATGAC GCTGAGTTAG TGGCGCAGCT GCGCTCTGAG
1931

20 AACGAGTTTT TACGGCGTCA GGTGAGCAG CAGGCGCGCA CGATCGAACC GCAGGCTGAG
1991

GCACACGCGG TGGTCTCAGC CCAGCTCACA CCGGTTGGCC AGCTTGAGGC CGGCGACGCA
2051

GCAGCACCGA CACTGGCACC CGTTGAAAGG CCGCCTCCGC GACGCGCGTG GTGGCAGCGT
2111

25 CGGTAGCGGT CAGGATCGCT CTGGCGTGAC GAGTGTGTCT GGCAGTCCGA ACAGTTGCTC
2171

GACCACTGGC AGCAGAAGCG AGATCGCTGC GTGGTGCTGT TCCTCGGTCA GTTCGTGAG
2231

30 GACTGGCGGG TCTTACTGCG TCCAGCCGAT CGCCTCGGCG GCCAAGGTCA GTTCCAAGCT
2291

GTGCCAAGCG ACACGCCCTT CCGCTGACAG CTCAGTCTCG AACTGTGCAA CTGGACCGGC
2351

CGGAAGATGC ACGTTGCCGA GGTGCTGAGT GGCCAAGCGC ACGTCAAAGA GTGCTGCTTC
2411

35 GTAGCCGCGC AGAAATGGCA GTGCTCGGTC GATTGAGATC GGCCTGCCCA GGTACATTCC
2471

GGGCCGCTTG ATGAACGCTT CCGCGTAGAA GCGCACCGTT CTCGGCCCCG CCTCGTGATC
2531

TGTCACTGTG CACGCTCCTC TCGATCCTTC TCGACGCTAC CGGAGACCAC CGACGTTTAT

09720583.062501

2591

CCCCAGCGCA GCGACCTGAA AGGACCAAGC CGAGTTAGCC GTGCTAACCG TATAGCTTGC
2651

5 TCCGTCGCCT CTGAGGGCAA CCACCTGCGC AGCAGGTGGG CGGCAGCCCC CGCGCAACCG
2711

CCTACCGGGT TTGGGCACAG CCCATAATC AACGCCTCCG GTGTTGAAGC GATCGTGTGT
2771

CACGATTGCT ATGCTTGCTA CCCCTTCAGG GTTTTCGTAT ACACAAATCA AGTTTTTTTCG
2831

10 TATACGCTAA TGCCATGAGT GAGCATCTAC TGCACGGCAA CCCCCTCACC AACGAGCAGA
2891

TTCAGGCATG GGCAGACGAG GCCGAGGCCG GATACGACCT GCCCAAATC CCCAAGCCAC
2951

15 GCGCGGAGCG CCCGCCCGTA GGAGACGGTC CGGGCACCGT CGTACCCGTG CGTCTCGACG
3011

CGGCCACCGT TGCCGCTCTC ACAGAACGAG CAACAGCCGA GGGCATCAG AACCGTTACG
3071

ACGCGATCCG AGCCGCGATC CACGAGTGGG CACCGGTTGC CTGACCTCCA CGACTCAGCA
3131

20 CGCAAGCACT ACCAACGAGA CCGGCTCGAC GACACGGCCG TCGTCTACGC GCGCACCCAC
3191

GTTCTCAACT CCCGGCCACT CGACGACGAA GACGACCCGC GCCGCTGGCT CATGATCGGA
3251

25 ACCGACCCAG CAGCCCGCCT ACTCGAACTC GTCGCACTGA TCTACGACGA CGGCTACGAA
3311

CTGATCATCC ACGCAATGAA AGCCCCCACC CAATACCTCG ACCAGCTCTA ACCAAGAAAG
3371

GAACCTGATG AGCGACCAGC TAGACAGCGA CCGCAACTAC GACCCGATGA TCTTCGACGT
3431

30 CATGCGCGAG ACCGCGAACC GCGTCGTCGC CACGTACGTT GCATGGGAAG ATGAAGCCGC
3491

TGATCCCCCG CAGGCTGCCG ACTGGCAGGC CGAGCGATTC CGCACCCGCG ACGAGGTGCG
3551

35 CGCC
3595

(2) INFORMATION FOR SEQ ID NO: 2:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 303 amino acids

09720589.062501
105290.6850260

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

5 Met Asp Ser Phe Glu Thr Leu Phe Pro Glu Ser Trp Leu Pro Arg Lys
1 5 10 15

Pro Leu Ala Ser Ala Glu Lys Ser Gly Ala Tyr Arg His Val Thr Arg
20 25 30

10 Gln Arg Ala Leu Glu Leu Pro Tyr Ile Glu Ala Asn Pro Leu Val Met
35 40 45

Gln Ser Leu Val Ile Thr Asp Arg Asp Ala Ser Asp Ala Asp Trp Ala
50 55 60

Ala Asp Leu Ala Gly Leu Pro Ser Pro Ser Tyr Val Ser Met Asn Arg
65 70 75 80

15 Val Thr Thr Thr Gly His Ile Val Tyr Ala Leu Lys Asn Pro Val Cys
85 90 95

Leu Thr Asp Ala Ala Arg Arg Arg Pro Ile Asn Leu Leu Ala Arg Val
100 105 110

20 Glu Gln Gly Leu Cys Asp Val Leu Gly Gly Asp Ala Ser Tyr Gly His
115 120 125

Arg Ile Thr Lys Asn Pro Leu Ser Thr Ala His Ala Thr Leu Trp Gly
130 135 140

Pro Ala Asp Ala Leu Tyr Glu Leu Arg Ala Leu Ala His Thr Leu Asp
145 150 155 160

25 Glu Ile His Ala Leu Pro Glu Ala Gly Asn Pro Arg Arg Asn Val Thr
165 170 175

Arg Ser Thr Val Gly Arg Asn Val Thr Leu Phe Asp Thr Thr Arg Met
180 185 190

30 Trp Ala Tyr Arg Ala Val Arg His Ser Trp Gly Gly Pro Val Ala Glu
195 200 205

Trp Glu His Thr Val Phe Glu His Ile His Leu Leu Asn Glu Thr Ile
210 215 220

Ile Ala Asp Glu Phe Ala Thr Gly Pro Leu Gly Leu Asn Glu Leu Lys
225 230 235 240

35 His Leu Ser Arg Ser Ile Ser Arg Trp Val Trp Arg Asn Phe Thr Pro
245 250 255

Glu Thr Phe Arg Ala Arg Gln Lys Ala Ile Ser Leu Arg Gly Ala Ser
260 265 270

T05290" E0502460

Lys Gly Gly Lys Glu Gly Gly His Lys Gly Gly Ile Ala Ser Gly Ala
 275 280 285
 Ser Arg Arg Ala His Thr Arg Gln Gln Phe Leu Glu Gly Leu Ser
 290 295 300

5 (2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 85 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Met Thr Thr Arg Glu Arg Leu Pro Arg Asn Gly Tyr Ser Ile Ala Ala
 1 5 10 15
 Ala Ala Lys Lys Leu Gly Val Ser Glu Ser Thr Val Lys Arg Trp Thr
 20 25 30
 Ser Glu Pro Arg Glu Glu Phe Val Ala Arg Val Ala Ala Arg His Ala
 35 40 45
 Arg Ile Arg Glu Leu Arg Ser Glu Gly Gln Ser Met Arg Ala Ile Ala
 50 55 60
 Ala Glu Val Gly Val Ser Val Gly Thr Val His Tyr Ala Leu Asn Lys
 65 70 75 80
 Asn Arg Thr Asp Ala
 85

(2) INFORMATION FOR SEQ ID NO: 4:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 59 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

AATCAAGCT TGTGACGTT AACCTGCAGG CATGCGGATC CGGTACCCAT ATCAGATCT
 59

(2) INFORMATION FOR SEQ ID NO: 5:

- 35 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 59 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CCGAAGATCT GATATCGGTA CCGGATCCGC ATGCCTGCAG GTTAACGTCC ACAAGCTTG
59

(2) INFORMATION FOR SEQ ID NO: 6:

- 10 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GTACCGGCCG CTGCGGCCAA GCTT
24

(2) INFORMATION FOR SEQ ID NO: 7:

- 20 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GATCAAGCTT GCGCGCAGCG GCGG
24

(2) INFORMATION FOR SEQ ID NO: 8:

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: DNA (synthetic)

FOI 99-06222-150

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

AAACTGCAGC TCCTGGCTTG CCCCCGATGC TAGTC
35

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 76 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

AAACTGCAGC AGCTGGGCAG GCCCCTGGAC GCCCTGCCCT CGAGCTCGTC TAGAATGTGC
60

TGCCGATCCT GGTTC
76

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CTAGTCTAGA CACCGATGAG GAAACCCGAT GA
32

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CCCAAGCTTC TCGAGTCACT GGTGCTGGG CCGCGG

105290" E8502760

36

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

10 GGAGATCTAGATCGATATCTCGAG
24

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

20 GATCCTCGAGATATCGATCTAGATCTCCGC
30

T.03290" E.0502/50